Appendix II: Lower Willamette River Sediment Load

Analysis of Sediment Loading: Motivation

Deposition and erosion patterns predicted by numerical sediment transport models typically show high sensitivity to sediment loading. For example, ±20% changes in tributary sediment loading to the Lower Passaic River produced changes in the range of ±25-125% in predicted deposition/erosion in some parts of that system [EPA Region 2, 2014]. Accordingly, time-dependent sediment inputs need to be specified accurately, sediment load rating curves are often used to for this purpose, and the Lower Willamette Group (LWG) modeling of Portland Harbor uses sediment load rating curves in Appencis La of the RI-FS and elsewhere. Yet rating curves are typically not accurate to better than 50% on an annual average basis [Gray and Simões, 2008; Meade et al., 1990] and may be highly inaccurate for individual events. It is, therefore, vital to analyze available Willamette River sediment loading data to determine the best representation of these data in the form of a rating curve, and to determine whether the LWG rating curves provide an accurate representation of sediment input.

Available Data

No sediment load or concentration data are available at the boundaries of the lower Willamette River at Oregon City and the mouth of the Clackamas River. Therefore, data from Portland Harbor have been assumed in the RI-FS to be representative of inputs from the upstream boundaries. Most data sets after 1960 for the Willamette River were collected at Portland (at Morrison Street Bridge); they include the following:

1. Daily US Geological Survey (USGS) discharge, sediment load and concentration data collected from July 1962 to September 1964 (http://co.water.usgs.gov/sediment/seddatabase.cfm). Sand discharge is presented graphically as a time series in Haushild et al. [1966], from which percent sand or fines can be derived. This is the only long time series of sediment discharge available for the system. Given the importance of sediment load hysteresis, it plays an important role in calibration of the sediment load. There are 822 daily samples.
2. Multiple (two to four) samples per day during the December 1964 flood, 20 December 1964 to 2 January 1965 (the last two days have only daily samples). Discharge, concentration and (in some cases) percent sand, silt and clay were determined, and loads calculated. These data were digitized from Waananen et al. [1970]. They are highly valuable because this was the largest Willamette River flood event of the last 60 years. Sediment load hysteresis was very prominent in this flood, and unusually well documented by the multiple samples each collected most days (Figure 1). Given the conditions, this was a remarkable sampling effort. There are 35 samples.
3. USGS NWIS sampling (1974 to 2014) focused primarily on water quality (http://nwis.waterdata.usgs.gov/or/nwis/qwdata). Discharge, sediment concentration and percent fines were typically measured, and discharge calculated[1]. Sampling is typically at monthly

[1] Discharge, given in tons/day, is calculated by a computer program that takes into account variability of discharge during the day. Thus, multiplying discharge by concentration does not always yield the stated sediment load [per-

intervals, with many exceptions, either missed samples or multiple samples per month. There are now 509 samples. In addition, to obtain information for the February 1996 flood, 5 samples taken at the St Johns Bridge were also included in the analysis. Unfortunately, there was no systematic sampling during this event, the second largest since 1950.

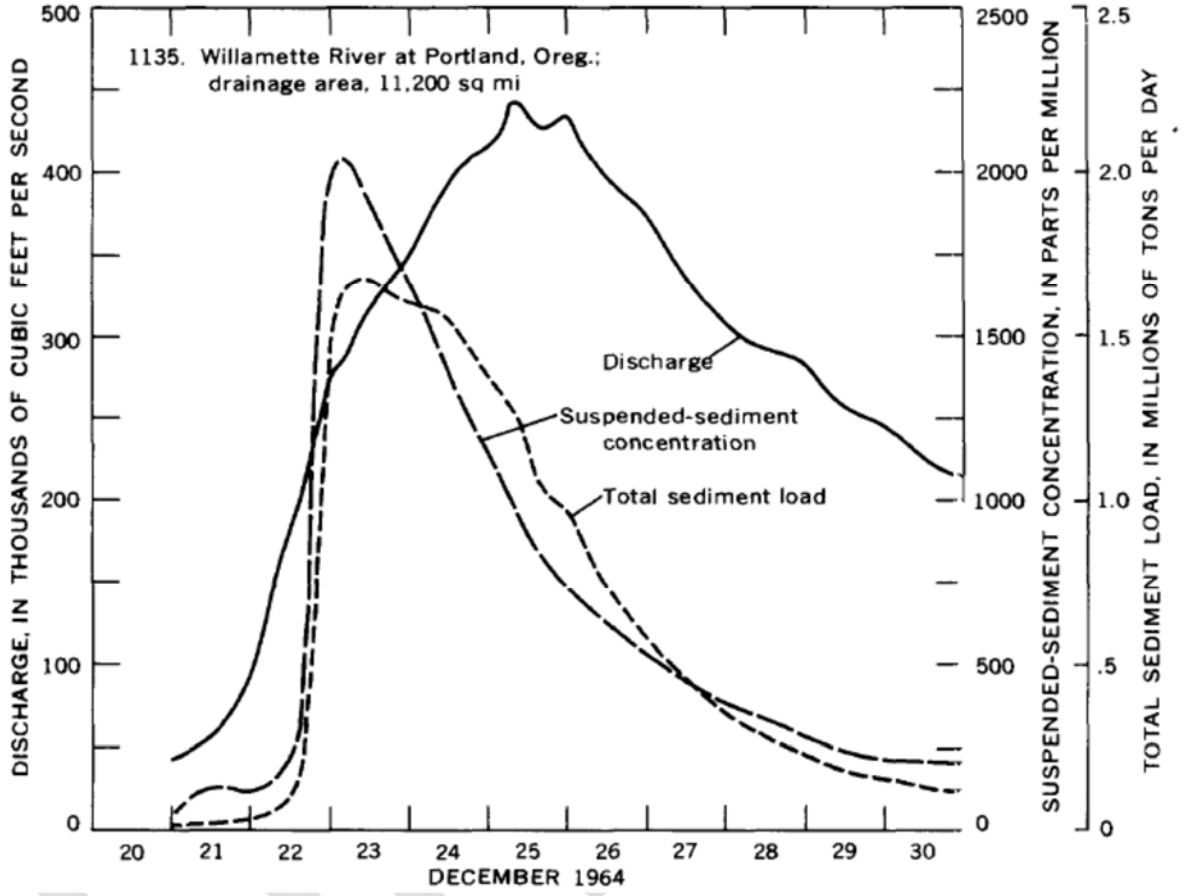

Figure 1: Discharge, suspended sediment concentration, and total sediment load during the flood of December 1964 from Waananen et al. [1971]. Note the strong hysteresis – maximum sediment concentration and load both occur on the rising arm of the hydrograph, before the peak of the flood. This is a typical pattern for the Willamette River. Sediment loads on the falling hydrograph are much smaller.

An additional data set has been collected further landward by the City of Portland, starting in the 1990s. In principle, these data might provide a better view of the sediment load from the Willamette and Clackamas Rivers, because the samples at the Morrison Street Bridge are affected by the load entrained or deposited within the system. However, these data were not available to the author, and suffer from methodological problems, in that they are not water-column integrated samples. There are additional data from earlier in the 20th century, but these were either not available or considered to be of historical interest only.

---

sonal communication 24 Nov 2014, D. Schoellhamer, USGS]. In principle, this methodological difference renders this data set different from the previous two. In practice, the three can be combined.

Preliminary examination of the data suggested that the above three data sets could be combined for analysis. The resulting total of 1367 total load, 867 fines load and 551 sand load samples is by no means excessive, given the complexity of the Willamette River sediment load. The number of sand samples is less than the number of fines samples, because samples that showed 100% fines were excluded from the sand load analysis.

Sediment Load Modeling for the Combined 1962-2014 Data Sets

Approach to sediment load modeling

The Willamette River at Portland sediment load observations exhibit two forms of complexity that make modeling the load complicated: a) there is a strong hysteresis effect, with the load being much higher on a rising hydrograph than a falling one (Figures 1 and 2); and b) the load vs. discharge exhibits a distinct change in slope at a level slightly above the mean discharge (Figures 3a,b). Figures 3a,b also emphasize: a) that the 1962-1964 and 1974-2014 data sets can reasonably be combined, despite the impacts of possible changes in land use; b) that the December 1964 data set is vital – all but one of the high-flow samples is from this event; c) that the sand load data are much more scattered than the fines or total load data; and d) from Figure 3a, that a transformation of the data set by a flow discharge ratio *R* (discussed in the next paragraph) reduces the scatter of the data by accounting for hysteresis.

Some experimentation with rating curve models was needed, given the observed hysteresis and change in slope exhibited by the data. The suggested non-dimensional total load rating curve is:

$$Q_{S,t} = aQ_t^b Max[\{1, \left(\frac{Q_t}{Q_C}\right)^c\}]{R_{t-L}}^d$$

$$R_{t-L} = \frac{Q_{t-L}}{Q_{t-L-1}}, \text{ Discharge ratio}$$

$$Q_{S,t} = \text{Load (non} - \text{dimensional) at time } t$$

$$Q_t = \text{Discharge (non} - \text{dimensional) at time } t$$

$$Q_C = \text{Critical discharge}$$

$$L = \text{Lag, days}$$

$$t = \text{Time, days}$$

$$a, b, c, d = \text{Regression coefficients} \qquad (1)$$

Discharge is non-dimensionalized by the long-term mean flow, 9786 m^3/s, while total load is non-dimensionalized by the median of the sediment loads for July 1962 to September 1964, 451 mtons/day (mtons = metric tons). A non-dimensional relationship is desirable so that the units of the constants do not vary with the exponents. This form of rating curve deals with the hysteresis effect via a lagged discharge ratio, $R_{t\text{-}L}$. Use of a threshold Q_C allows the change in slope to be accommodated within a single

regression analysis. If the low-flow and high-flow data are analyzed separately, then two problems arise: a) there is insufficient dynamic range of flow for low flows to achieve an accurate regression, and b) there are too few load observations to achieve an accurate regression for high flows. Therefore, a unified regression analysis that includes all samples, as performed here, is preferable.

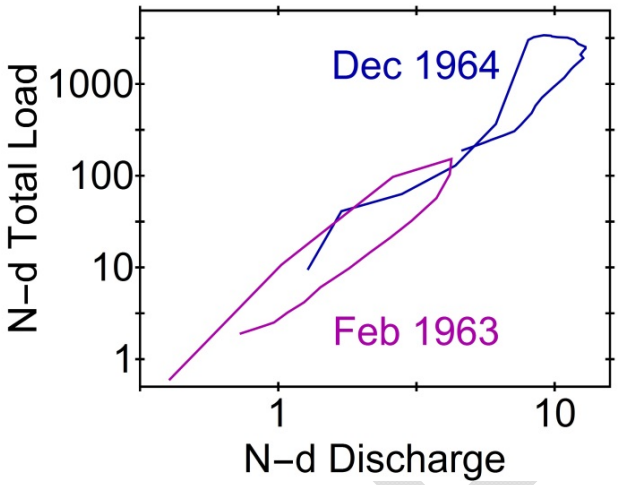

Figure 2: Examples of sediment load hysteresis for the Willamette River at Portland: scatter plot of non-dimensional total load vs. non-dimensional discharge, connected in chronological order. In both cases, the rising arm of the hydrograph shows higher load, by as much as an order of magnitude. Hysteresis occurs whenever the flow increases sharply, for high or low flows. The February 1963 example covers 17 days, while the December 1964 example (13 days) covers the 1964 flood, the largest flow since 1950.

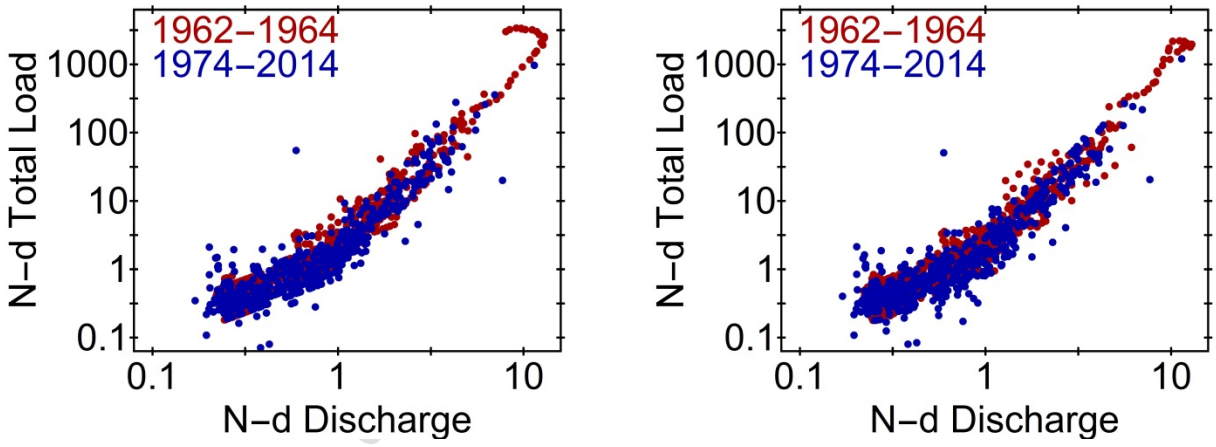

Figure 3a: The total load data set plotted in two forms: at left, the load is plotted in non-dimensional form; at right, the non-dimensional load is transformed by dividing by the hysteresis factor *R*; this removes the effect of hysteresis from the data set. Even after this transformation, there is still a change in slope at about the level of the mean discharge.

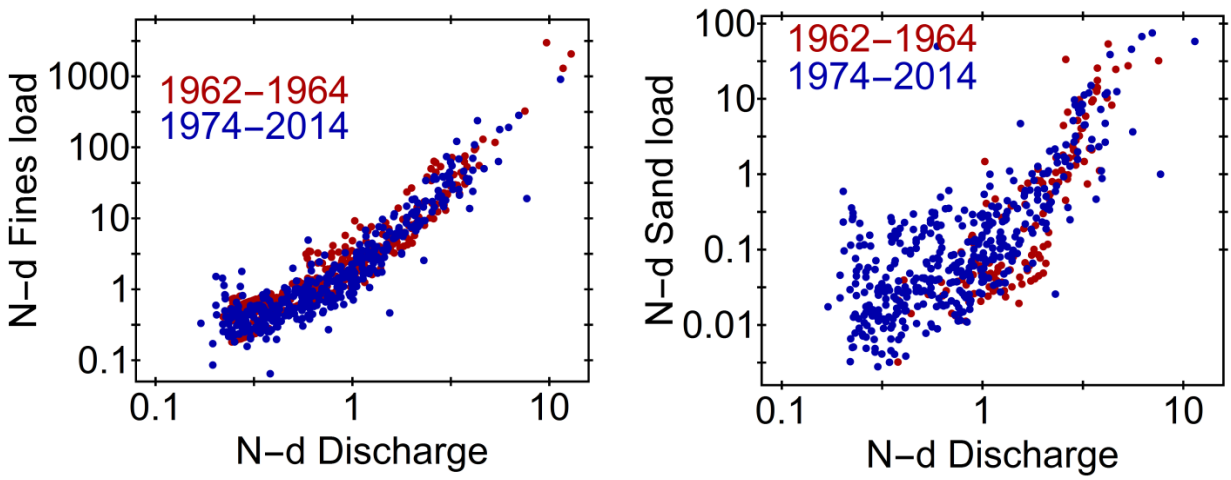

Figure 3b: The fines load data at left and the sand load data at right, both non-dimensional and plotted against non-dimensional flow. The non-dimensional loads have not been transformed, so both panels are comparable to Figure 3a (left panel). Note the large scatter in the sand load data.

The parameters *a*, *b*, *c*, and *d* in (1) are determined via a robust, multiple linear regression. A robust linear regression iteratively down-weights outlying points in a regression without totally removing them from the analysis [Leffler & Jay, 2009]. With scattered data, it achieves more accurate results and provides tighter confidence limits on model coefficients than a conventional regression. One more detail of the regression is important: the initial weighting of the data points. High flow days with high sediment loads are rare in a flashy river like the Willamette, and a much longer record than the 1367 available points (equivalent to only 3.74 years of daily data) is needed to capture a representative sample of the high flows that transport most of the sediment. This situation is dealt with by using a flexible weighting approach that sets the weight *wt*[*t*] of the data point at time *t* according to:

$$wt[t] \sim Log[10, Q_S(t)]^n \quad 0 \le n \le 6 \qquad (2)$$

The value of *wt*[t] has, however, a ceiling and is not allowed to exceed a chosen value of 1000-5000. The exponent *n* can be set in a variety of ways, depending on the character of the data. In this case, initial experimentation with the data suggested that the range $0 \le n \le 6$ was sufficient to provide realistic regression model results. Values of $n > 6$ weighted high flows excessively and resulted in unrealistic results. The final value of *n* was chosen by performing the robust regression for values of *n* =0, 0.2, 0.4, …, 6, and choosing the "best" result. Result quality was determined by the following factors: a) maximizing the adjusted R^2 of the analysis; b) correctly reproducing the mean transport; and c) reducing the standard deviation and bias of analysis residuals. The lag *L* was determined, based on the quality of the results, for *L* = 0, 1, and 2 days. While non-integer lags might be useful with more detailed discharge information, this was not practical with the existing (daily) USGS discharge data. The critical nondimensional discharge Q_C was chosen as $\text{Log}_{10}[Q_C]$=0.1 after testing of $\text{Log}_{10}[Q_C]$=0.0, 0.1, 0.2 and 0.3. Note that for $\text{Log}_{10}[Q_C]$=0.1, Q_C =1.259; i.e., ~126% of the mean discharge.

Results for total load modeling

Results of the regression analysis described in the previous section are shown in Figure 4. The adjusted R^2 of the model is 0.985, and the residuals are well balanced between positive and negative. Clearly, the model tracks the overall pattern of the data, but isolated large residuals remain, and relative errors can be large for low flows. One important point in interpretation of the low-flow parts of log-log plots of Figure 4a and b: the curve is above most of the data points, which is appropriate for an unbiased fit to data displayed on a log-log curve. This point arises from the fact that Log[Average[{a_1, a_2}]] > Average[Log[{a_1, a_2}]] for any two number a_1 and a_2, $a_1 \neq a_2$. Thus, a regression fit to scattered data in log-log space will normally result in a model that under-predicts sediment load, and the under-prediction increases as the scatter increases. This can be dealt with by what is called smearing corrections. Here, the weighting used results in an unbiased load estimate without resorting to smearing corrections (Duan, 1983). The appropriateness of the model shown in Figure 4 is further tested below, by comparison with data for the 1964 and 1996 floods. An optimal lag of L =1 day was found. This is consistent with the idea that the total load is influenced by the supply of the material from upstream, and/or that the sediment load moves more slowly than the flow, especially in the tidal part of the system. Table 1 summarizes model parameters for all load models.

Table 1: Sediment Load Regression Model Parameters

	Parameter	**L**	**$Log_{10}[Q_C]$**	**a**	**b**	**c**	**d**	**n**
Model								
Total load		1	0.1	0.4212±.0174	1.2202±.1321	1.4797±.1405	1.3178±.0362	5.4
Fines load		1	0.1	0.3573±.0161	1.0919±.0866	1.5089±.1048	1.3845±.0736	4.4
Sand load		0	0.1	-0.9373±.0442	0.6008±.1685	2.966±.2692	1.3600±.4082	1.6

Approach and results for fines load modeling

The multiple linear regression load model described by (1) and (2) was also applied to the fines load data shown in Figure 3b (left panel). As before, $Log_{10}[Q_C]$ =0.1. Results are shown in Figure 5. The adjusted R^2 of the model is 0.979. This is slightly lower than for the total load model and may reflect the smaller number of data points. The residuals are reasonably well balanced between positive and negative, but smaller in magnitude than for the total load. The model tracks the pattern of the data quite well, similar to the total load. An optimal lag of L =1 day was found, as for the total load. This is consistent with the fact that fines make up most of the load under most conditions.

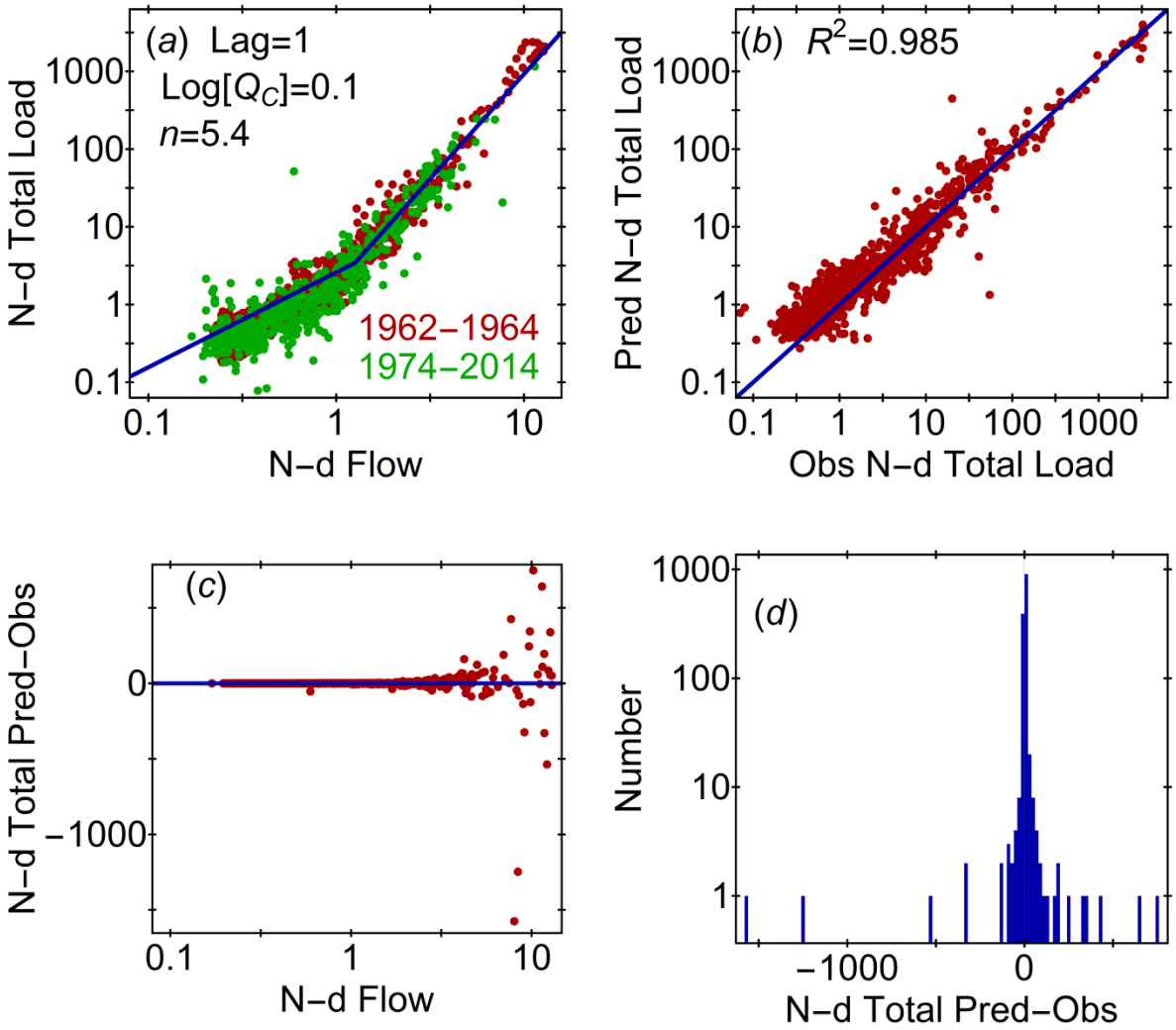

Figure 4: Total load regression model results: a) non-dimensional total load vs. non-dimensional flow, for the data (red 1962-1964 and green 1974-2014 points) and model (blue line) from (1) and (2); b) model predicted vs. observed non-dimensional total load with a 1:1 line (blue); c) model residuals vs. non-dimensional flow; and d) a histogram of residuals. In a), the model line is plotted with the median value of *R* (i.e., Log[R]= -0.004); thus, the data are transformed as in Figure 3b.

Approach and results for sand load modeling

The multiple linear regression model described by (1) and (2) was also applied to the sand load data shown in Figure 3b (right panel). As before, $\text{Log}_{10}[Q_C]$ =0.1 but L =0. Results are shown in Figure 6. The adjusted R^2 of the model is 0.858. This is lower than for the total load and fines load models; this may reflect the fact that the number of data points is much smaller than for either total load or fines load. The residuals are smaller than for the total load, but rather skewed. The model tracks the pattern of the data quite well for flows above the mean flow, but the observations are very scattered at low flow, and the predictions for low flows do not drop as low as some of the observations. One reason for the scatter

at low flow levels is the fact that the sand load is determined by difference and accumulates errors. The fact that lag L=0 worked best for the sand model is consistent with the idea that the sand load is transport capacity limited and responds to local conditions.

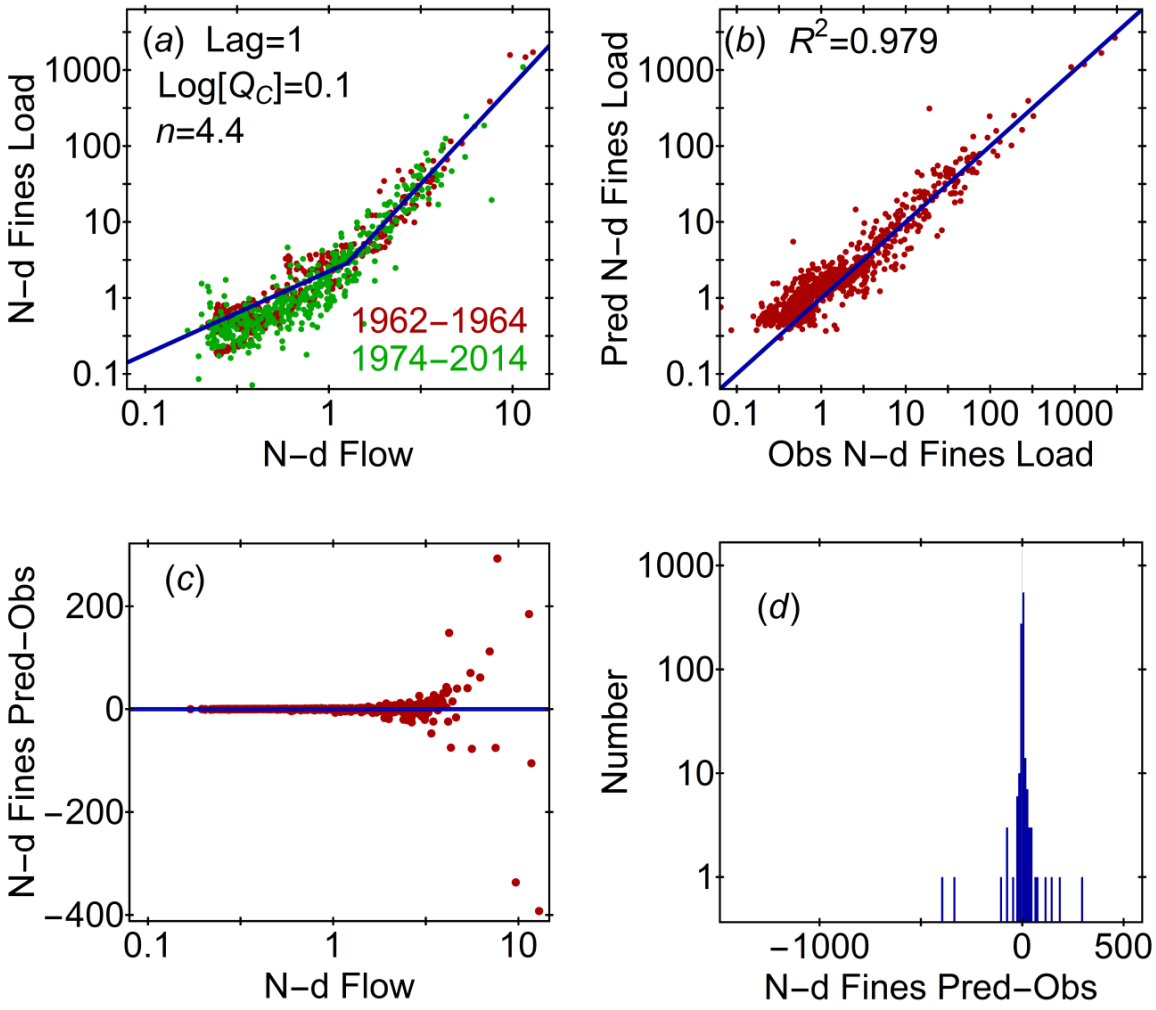

Figure 5: Fines load regression model results: a) non-dimensional fines load vs. non-dimensional flow, for the data (red 1962-1964 and green 1974-2014 points) and model (blue line) from (1) and (2); b) model predicted vs. observed non-dimensional fines load with a 1:1 line (blue); c) model residuals vs. non-dimensional flow; and d) a histogram of residuals. In a), the model line is plotted with the median value of *R* (i.e., Log[R]= -0.004); thus, the data are transformed as in Figure 3b.

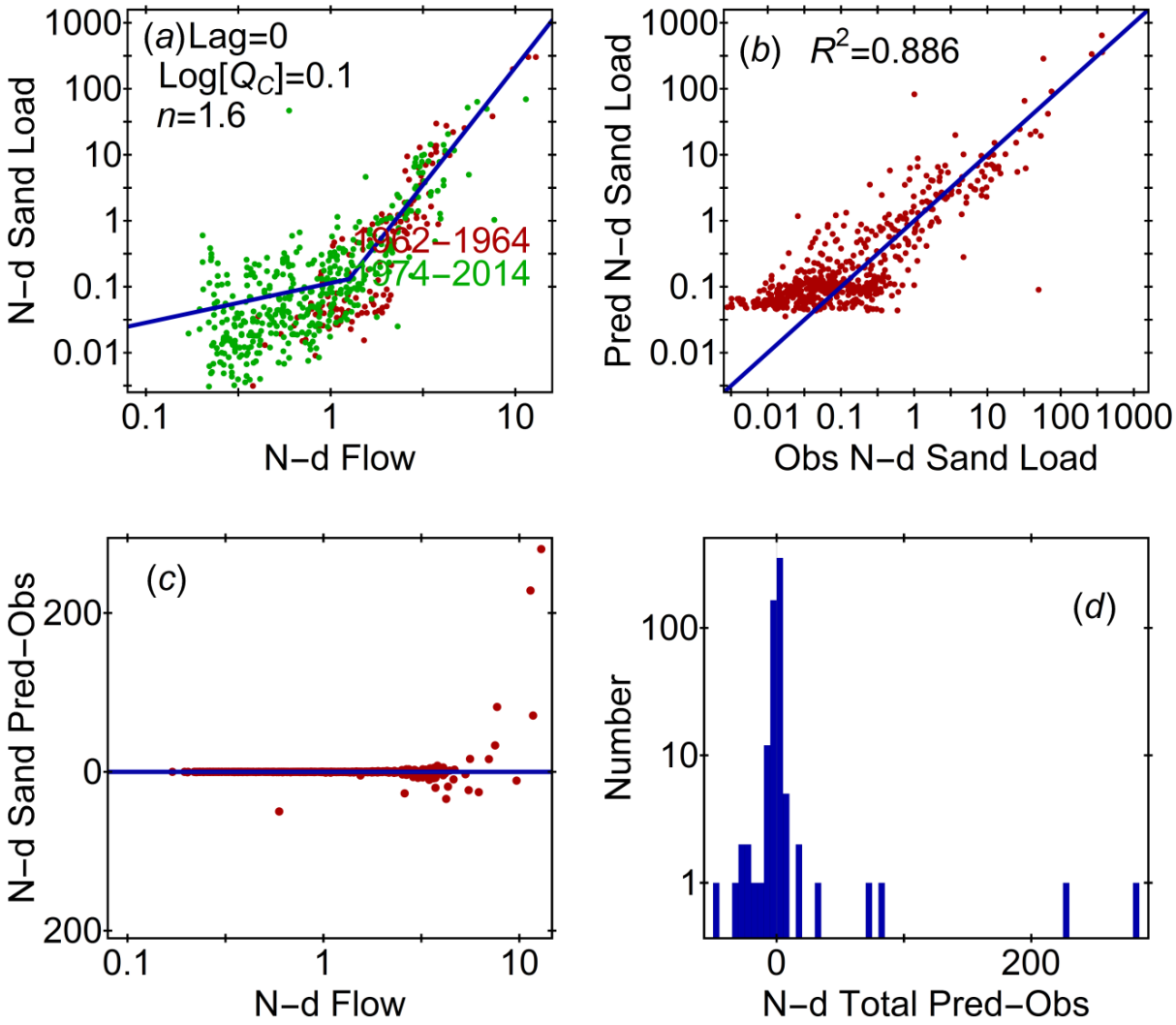

Figure 6: Sand load regression model results: a) non-dimensional sand load vs. non-dimensional flow, for the data (red 1962-1964 and green 1974-2014 points) and model (blue line) from (1) and (2); b) model predicted vs. observed non-dimensional sand load with a 1:1 line (blue); c) model residuals vs. non-dimensional flow; and d) a histogram of residuals. In a), the model line is plotted with the median value of *R* (i.e., Log[R]= -0.004); thus, the data are transformed as in Figure 3b.

Consistency of the load models

It is important that the independently derived models of the total, fines and sand loads be consistent, i.e., that the sand and fines load sum (at least approximately) to the total load. This is verified in Figure 7. The sum of the fines plus sand loads is very close to the total load (within the confidence limits on the total load line), except at non-dimensional flows <0.5, which corresponds to flows less than ~400m^3/s. It is also noticeable that the confidence limits on the sand load are much broader at high flows than those for the fines and total load. Given that sand transport is transport-capacity limited, this is a reflection on the small sand transport data set, and the fact that sand transport values were derived by difference.

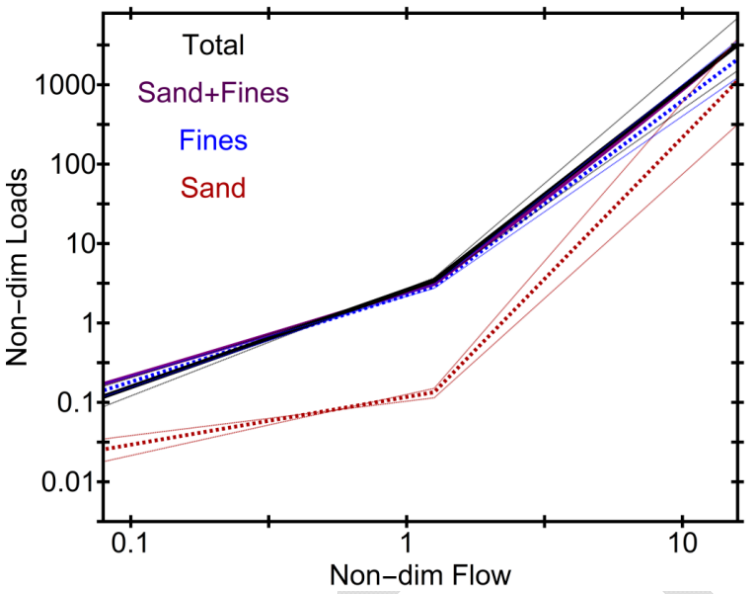

Figure 7: Rating curves for the total, fines and sand loads, and the sum of the sand and fines loads. Error bars are shown, based on 95% confidence limits of all parameters. The purple sand+fines load is under the total load line for most values of flow.

Finally, comparing the error bars in Figure 7 with the scatter of the data in Figures 3a,b it appears that the failure of models to fully capture the variability of the data is not merely a question of statistical parameter uncertainty, but a reflection of the inability of a simple model like that provided by (1) and (2) to capture the full variability of the complex sediment load processes affecting the lower Willamette River and Portland Harbor.

In summary, (1) and (2) provide a framework that allows consistent and reasonably accurate modeling of Willamette River sediment load (total load, fines and sand) at Portland (Morrison Street Bridge). We do not know, however, how closely the load at this location matches the load input to the system by the Willamette and Clackamas Rivers.

Lower Willamette Group Sediment Load Modeling

According to LWG [2006], rating curves were developed during Phase 1 for total load and fines (or cohesive) load (Figures 8a,b), with separate calibrations for high and low flows. These relationships were used in lieu of relationships developed in Phase 2, because they appear to be more accurate. For the total load the relationships are:

$$Q_S = 0.1742\ Q^{0.7927} \qquad Q \leq 27{,}000 \text{ CFS}$$

$$Q_S = 7.31\times10^{-9}\ Q^{2.4605} \qquad Q > 27{,}000 \text{ CFS} \qquad (3a,b)$$

For the fines load, the relationships are:

$$Q_{sf} = 0.0987\ Q^{0.8499} \qquad Q \leq 28{,}000 \text{ CFS}$$

$$Q_{sf} = 7.85\times10^{-9}\ Q^{2.4433} \qquad Q > 28{,}000 \text{ CFS} \qquad (4a,b)$$

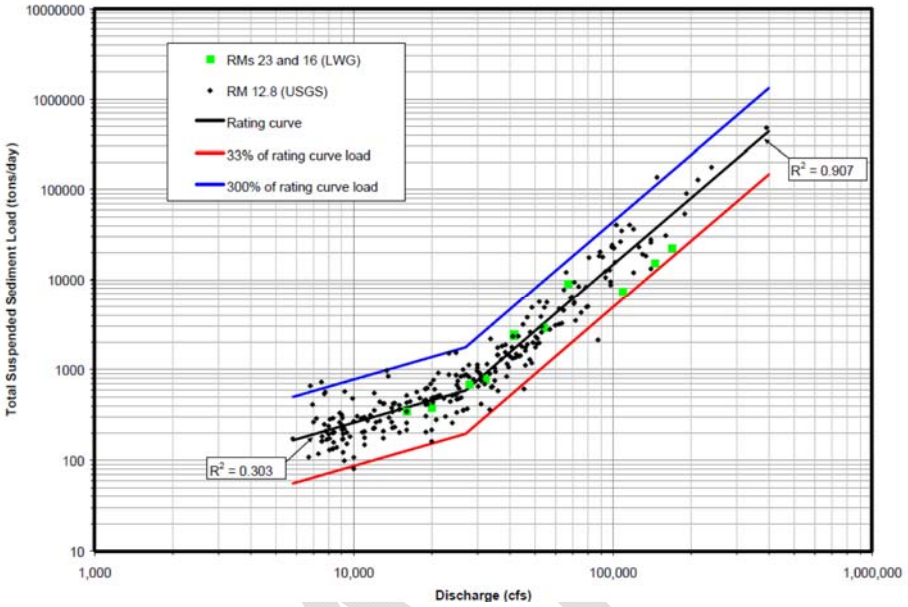

Figure 8a: Phase 1 total load rating curve; (Figure 2.7 in LWG [2006]). Dimensional units are used, and the loads are in short (US) tons/day.

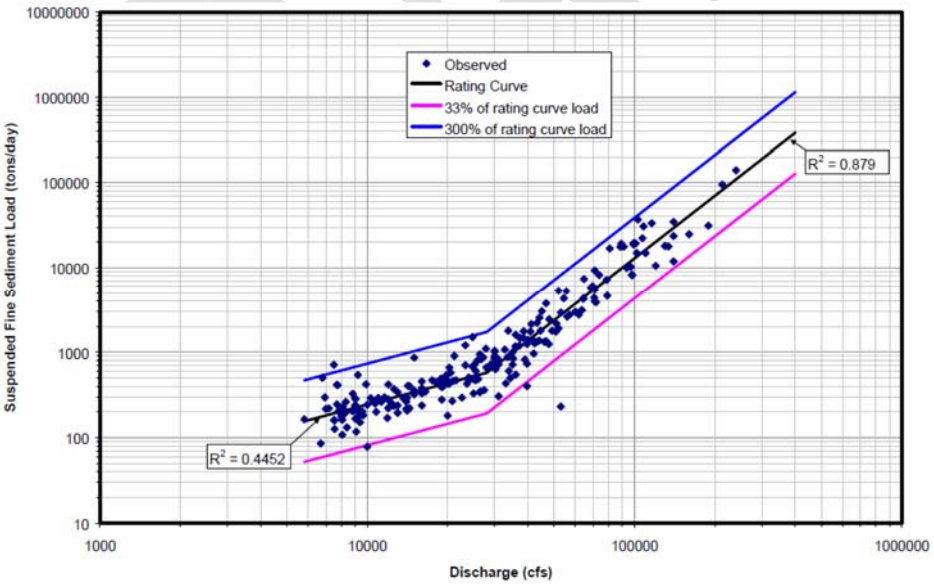

Figure 8b: Phase 1 fines rating curve; (Figure 2.8 in LWG [2006]). Dimensional units are used, and the loads are in short (US) tons/day.

These are dimensional relationships, with discharge in CFS (cubic feet per second) and sediment load in short (US) tons. Note that the exponent for the fines load for flows ≤28,000 CFS is greater than the exponent for total load for Q ≤27,000 CFS. This creates the possibility that the fines load may exceed the

total load for some flow levels. Note also that the R^2 values shown in Figures 8a,b are lower than those in Figures 4 and 5, especially for the low-flow part of the model. There are several likely reasons why the models are less successful. For low flows, the dynamic range of the independent variable (river discharge) is only about a factor of 2 and the range of mean sediment loads is <4×, while the scatter in the sediment load at any given flow level is as much as a factor of 9. Also, the simplicity of the model precludes consideration of hysteresis, an important factor for all flow levels. It is unclear whether smearing corrections were made, as the text does not clarify this point.

The models of (3) and (4) were superseded in Appendix La of the RI-FS by the following relationship for total suspended sediment concentration SedCon in mg/l (Figure 9):

$$\text{SedCon} = 9 \qquad Q \leq 33{,}000 \text{ CFS}$$

$$\text{SedCon} = 1.3 \times \left(\frac{Q}{10^4}\right)^{1.65} \qquad Q > 33{,}000 \text{ CFS} \qquad \text{(5a,b)}$$

These relationships are equivalent to total load model (in tons/d) of:

$$Q_S = 0.00269 \times 9\, Q \qquad Q \leq 33{,}000 \text{ CFS}$$

$$Q_S = 1.3 \times 0.00269 \times \left(\frac{1}{10^4}\right)^{1.65} Q^{2.65} \qquad Q > 33{,}000 \text{ CFS} \qquad \text{(6a,b)}$$

Note that these are short tons, not metric tons (mtons). The fines sediment load was then defined in terms of the fines fraction:

$$\text{FinesFraction} = 0.92 \qquad Q \leq 66{,}400 \text{ CFS}$$

$$\text{FinesFraction} = 1.01 - 0.014 \times \left(\frac{Q}{10^4}\right) \qquad Q > 66{,}400 \text{ CFS} \qquad \text{(7a,b)}$$

Figure 10 compares the LWG total load models of (3a,b) and (5a,b) to the 1962 to 2014 data. It is evident that the LWG models under-predict the high-flow loads that occurred during the December 1964 flood, leading to very large negative residuals. Even without the flood period, however, the model under-predicts loads, for reasons that are not clear. Similar remarks apply to the fines load model (Figure 11).

According to WLG [2006], the Phase 1 sand load was determined as the difference between the total load predicted by (3a,b) and the fines load from (4a,b). The Appendix La sand load was also derived by difference. It is important, therefore, to consider the consistency and accuracy of this formulation. Figure 12 shows LWG predicted sand load as the difference between total load and fines load. It is evident that the predicted Phase 1 fines load exceeded the predicted total load for a range of N-d discharge levels of ~0.6 to 0.9. leading to negative sand loads. The Appendix LA model corrects this particular problem. Nonetheless, Figures 10-12 suggest that the Appendix LA models predict loads that are too small.

In summary, the LWG load models of (3a,b) and (7a,b) do not provide a particularly good description of the sediment load of the Willamette River at Portland (Morrison Street Bridge). Flood loads are underes-

timated, a problem that is most acute for the fines load. The Appendix LA sand load appears to be closer to the data, but this is because the fines load has been underestimated. If the fines load were accurate, the sand load would be too small – this is a necessary consequence of the difference approach used for the sand load, and the underestimation of the total load.

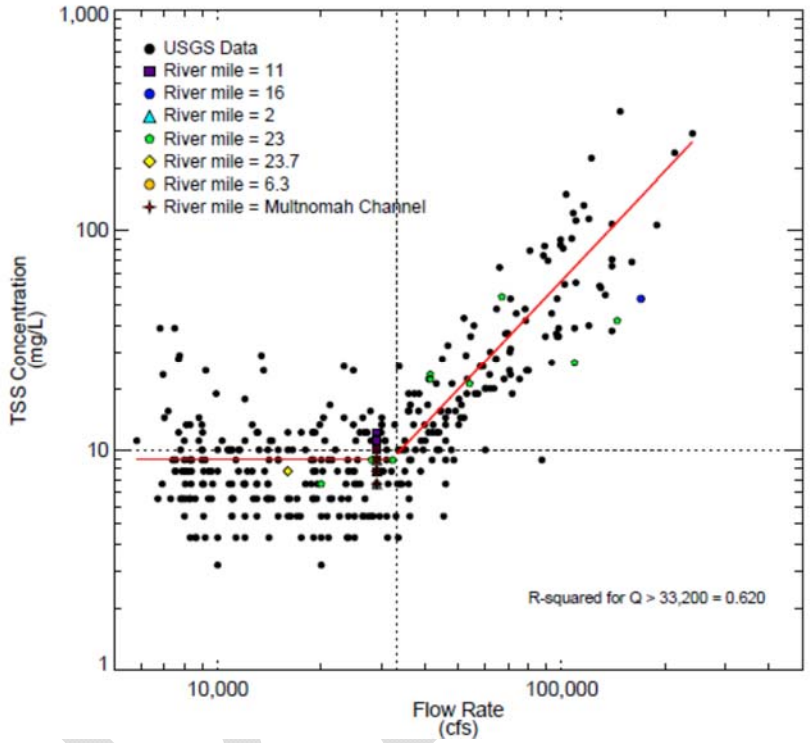

Figure 9: The LWG model of suspended sediment concentration from (5a,b); from Appendix La of the RI/FS.

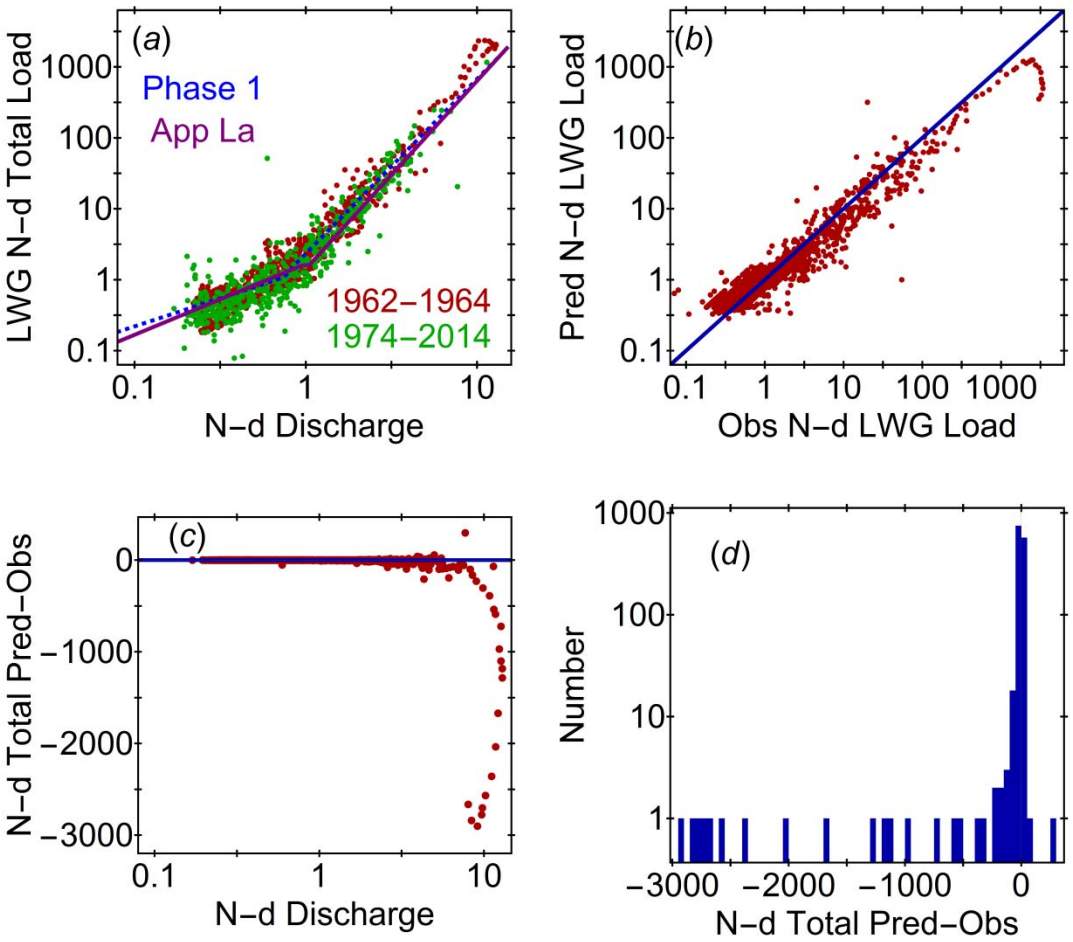

Figure 10: Comparison of the LWG Phase 1 total load model (3a,b) and Appendix La total load model (5a,b) to the 1962-2014 data; panels are as in Figure 4, except that the observations in a) were not transformed to remove hysteresis. Hysteresis was not removed, because the LWG model does not consider hysteresis. Also, the predicted data in b) are from the Appendix La model.

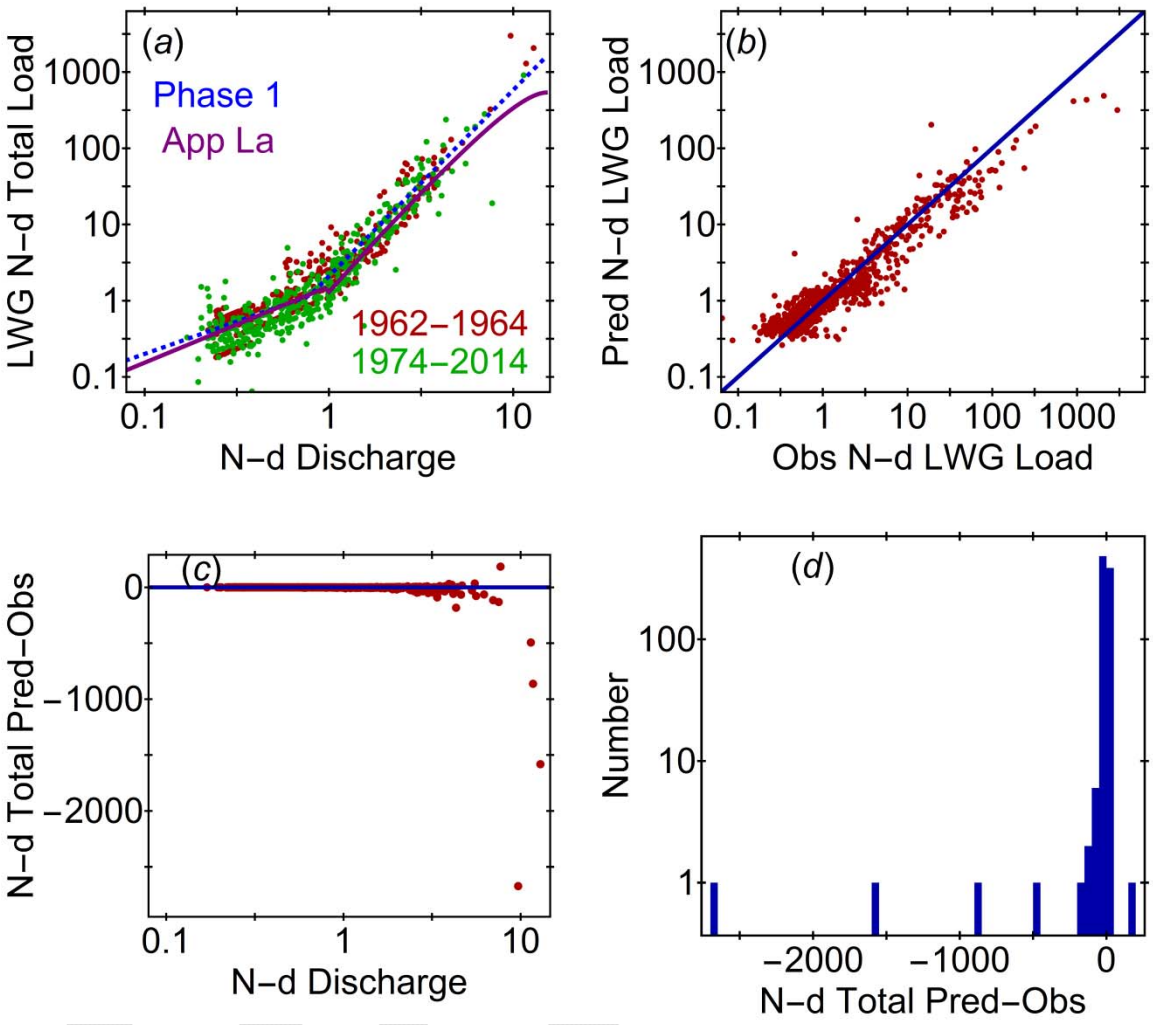

Figure 11: Comparison of the LWG Phase 1 fines load model of (4a,b) and the Appendix La model of (5a,b) to (7a,b) to the 1962-2014 data; panels are as in Figure 10.

To understand the problems with the Appendix La load models, it is useful to examine the fine fraction as a function of flow, using the full 1962-2014 data set (Figure 13). It is evident that the fines load does not continue to decrease as flow increases. The likely cause of this is that the high flows in the sediment record, and indeed most very high flows, are the result of rain-on-snow events. All of the largest events since 1923, i.e., those in 1943, 1955, 1964, 1974, 1996, and 1997 have had this character, with 1923 being, however, a major exception. Rain-on-snow events cause very high sediment loads because the rainfall cannot be absorbed, and in some areas, because frost heave has loosened the soil. The Willamette Valley has ample clayey soils to feed erosion, and clay loads are very high, particularly during the early stages of the flood, leading to at least part of the hysteresis that is so prominent in the Willamette River. It is an open question whether high flows and floods (like 1923) without extensive snowmelt exhibit similar patterns. In any event, a rating curve approach like (1) and (2), rather than a fractional load model appears to be a better approach for modeling the Willamette River fines load.

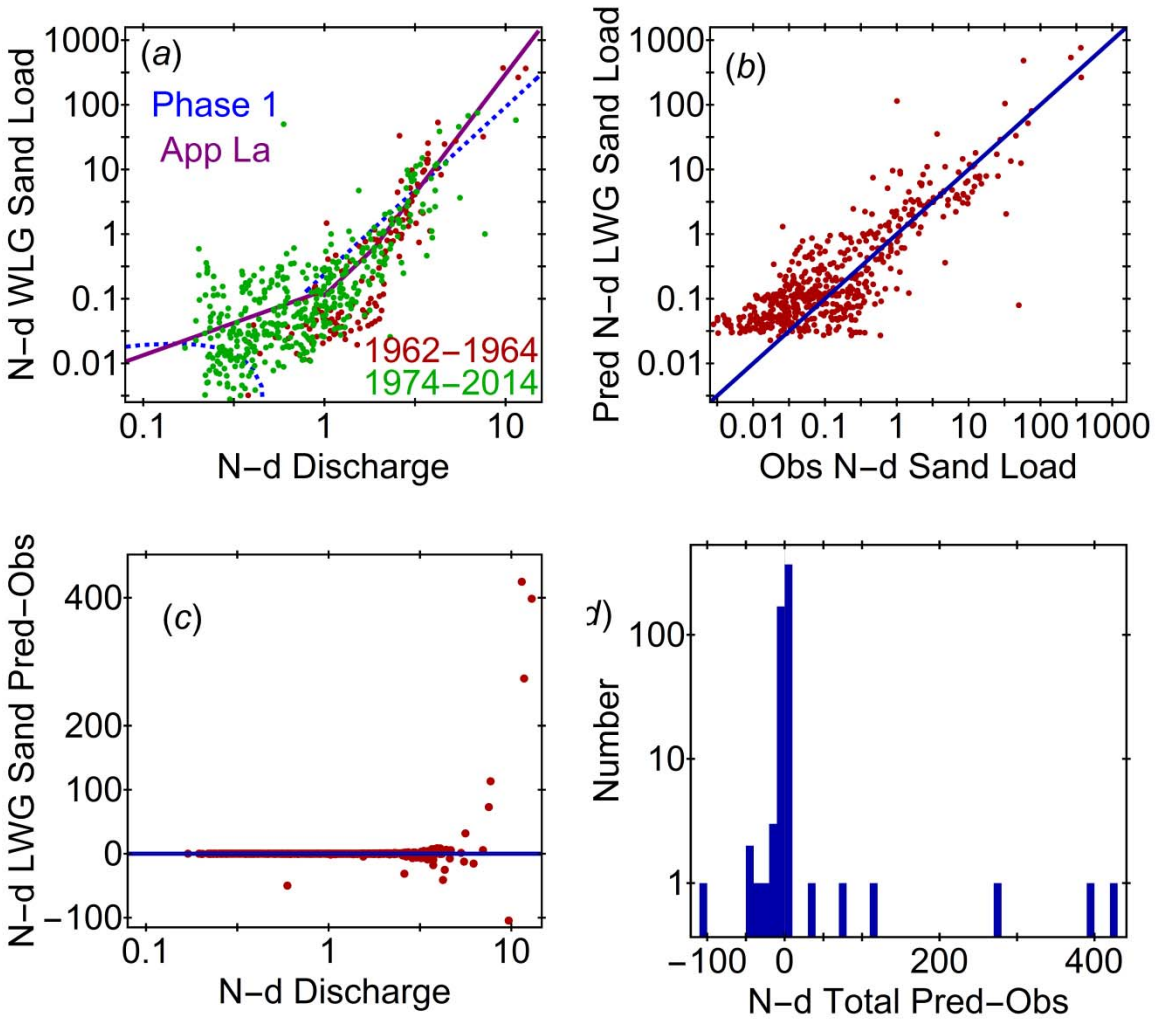

Figure 12: Comparison of the LWG Phase 1 sand load and the Appendix La sand load model implied to the 1962-2014 data; panels are as in Figure 10. The Phase 1 and Appendix La LWG sand load models are the difference between the total load and the fines load. The negative sand loads predicted in Phase 1 by this approach for N-d discharge levels of ~0.6 to 0.9 cannot be plotted on the log-log plot in a).

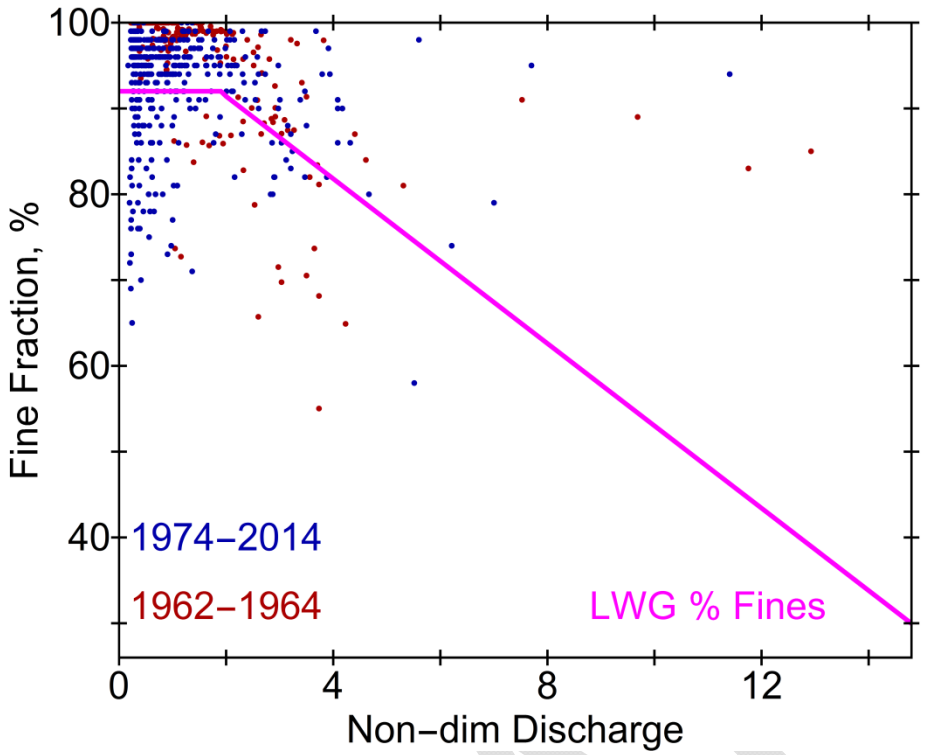

Figure 13: The Appendix La fines fraction relationship plotted against the available data. The highest non-dimensional discharge plotted corresponds to the 100-year flood (~500,000 cfs or 14,200 m^3/s). Though there are few data available for flows above about 200,000 cfs or 5700 m^3/s, it is evident that the fines fraction, variable as it is, does not decrease at high flows. It is lowest at a non-dimensional discharge of ~4, which corresponds to a dimensional discharge of about 3900 m^3/s. This may be a result of the fact that the highest flows for which sediment load data are available (and most of the historic high flows) have all been the result of rain-on-snow events that yield very high clay loads.

A Times-Series View of Willamette River Total Load

Three time periods have been selected to provide a time series view of the Willamette River sediment load and models thereof: a) winter 1962-1963, chosen because it was a winter with several moderate flow events, for which daily observations are available; b) the December 1964 flood, the largest such event in the last 60 years, for which a detailed data set is available; and c) the February 1996 flood, the second largest flood of the last 60 years, but with very few data.

Winter 1962-1963

Figure 14 shows discharge and total sediment load (observed and predicted) for the period from late November 1962 to mid-May 1963. The load data are daily, whereas the flow has been interpreted to 6-hr intervals, to provide a more detailed prediction of total load. The total load model from (1) and (2) generally tracks the load, but considerably over-predicts the highest peak load in early February 1963 and somewhat over-predicts the load in May 1963. The WLG model under-predicts most peak loads except May 1963. Both models appear to perform adequately at low flows. More specifically, the average

of the daily observed total load is 5275 mtons/day, with a maximum of 68,700 mtons/day. The mean and maximum predicted loads from the hysteresis model [i.e., based on (1) and (2)] are 6076 and 128,500 mtons/day, respectively. In contrast, the LWG total load model yields mean and maximum predicted loads based on (3) and (4) are 3408 and 23,450 mtons/day, respectively. The model based on (1) and (2) overstates the load for this time period, while the LWG model severely underestimates it. Favoring the hysteresis model is the point that the daily observations, probably collected as a standard time each day, may well have missed the peak load in February 1963. Removing one day of hysteresis model prediction at the peak of the February 1963 flood reduces the hysteresis model mean to 5430 mtons/day, very close to the correct result. Thus, almost the entire hysteresis model over-prediction for the period comes from this one day, emphasizing the difficulties in correctly modeling load. Nonetheless, it is apparent from this exercise that both total sediment load models yield inaccurate results even when averaged over almost half a year, and may give erroneous results for specific events.

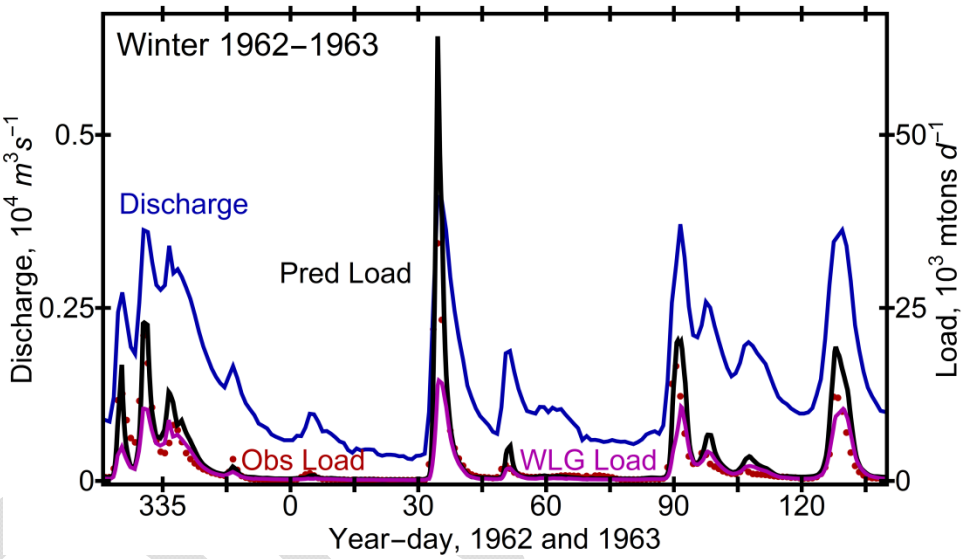

Figure 14: Observed and predicted total loads for the period from late November 1962 to mid-May 1963. See text for details.

The December 1964 flood

The December 1964 Willamette River flood was the largest of the last 60 years, though there were nine larger floods between 1861 and 1945. The sediment load for this event is also well documented. Thus, it is appropriate to examine model predictions for this period. Figures 15 and 16 show, respectively, the total load, and the fine and sand loads, for a 21-day period in December 1964 to January 1965. Figure 13 shows the total load observations plus daily and 6-hr model predictions. The LWG predictions of total load underestimate the peak load by a factor of ~4×. Moreover, the LWG total load is not delivered at the correct time, either in the daily or 6-hr predictions. This is because the LWG model does not account for hysteresis. The hysteresis-model 6 hr predictions of total load are accurately timed and represent the load for most of the event quite well. The peak total load is over-estimated, however, by about 10-12%. Significantly, the hysteresis model daily load predictions underestimate the peak and do not capture its

timing correctly. Thus, it is important during flood events to update the rapidly changing discharge and sediment load more often than daily.

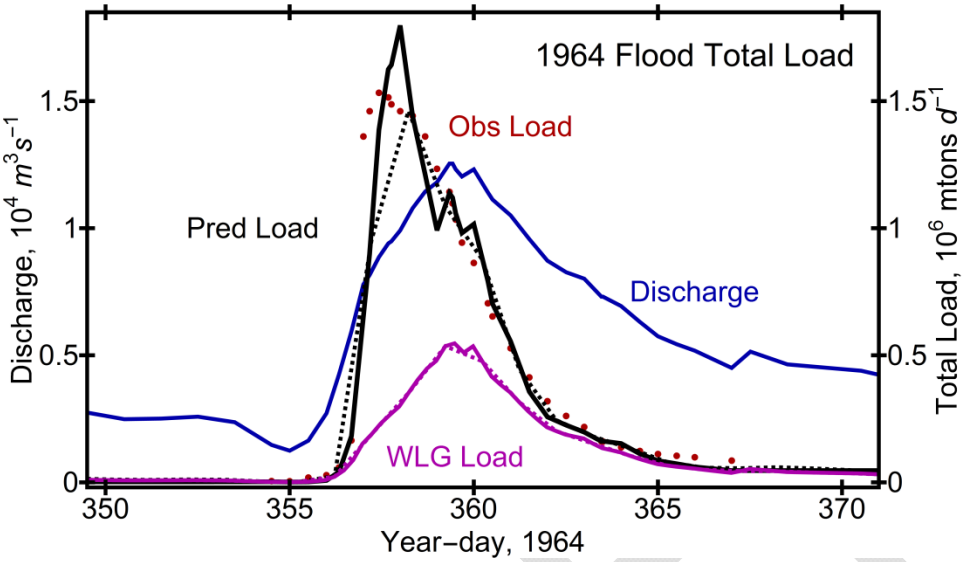

Figure 15: Observed and predicted total loads for the period of the December 1964 flood. See text for details. Note that the load is presented in 10^6 mtons/day instead of 10^3 mtons/day, as used in Figure 12. The dashed prediction lines correspond to daily predicted loads, while the solid prediction lines represent 6-hr predictions. Observations appear as dots.

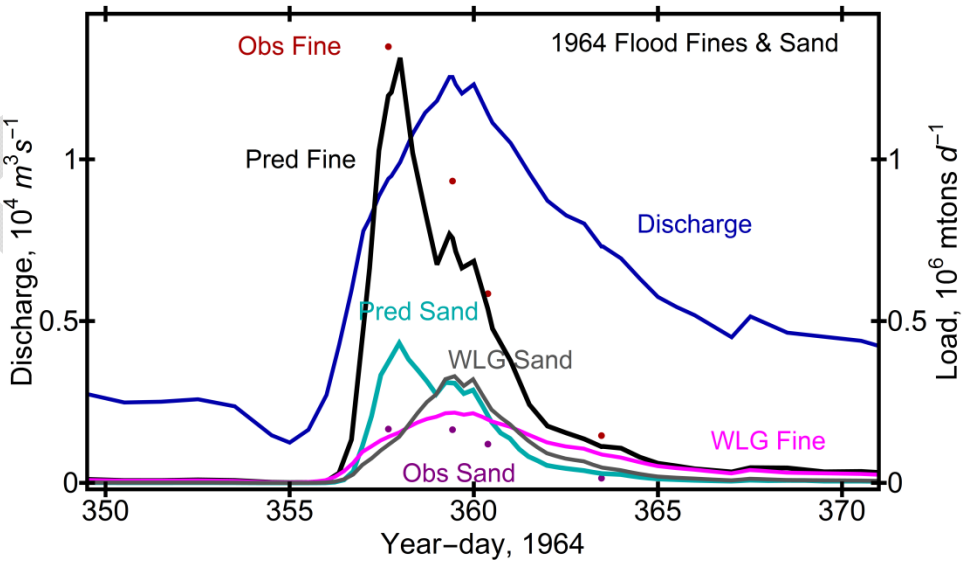

Figure 16: Observed and predicted fine and sand loads for the period of the December 1964 flood, based on 6-hr predictions. See text for details. The load is presented in 10^6 mtons/day but the vertical axis is different from Figure 15. Observations appear as dots.

There are only four observations of fine/sand load during the flood period. The hysteresis model represents the timing and volume of the fines load fairly well. The timing of the sand load predicted by the the hysteresis model is correct, but the volume is overstated by a factor of about two. The LWG model under-estimates the peak sand and fines discharge and timing of both the fine and sand transports are wrong; the peaks in both are too late and too small. That the hysteresis model correctly predicts the fines load but overestimates sand load emphasizes both the difficulty in modeling sand load and the statistical difficulties in measuring sand load indirectly, as the difference between the total and fines loads.

An overall impression of model performance is provided by estimation of the total load for the December 1964 flood period. USGS estimates that the sum of the daily total sediment loads for 21 to 31 December 1964 was 5.94×10^6 mtons [Waananen et al., 1970], an average of 541.6×10^3 mtons/day over 11 days. The hysteresis model estimates 5.56×10^6 mtons total load for the 11-day period, or 505.5×10^3 mtons/day, 93.3% of the USGS estimate. The LWG total load model estimates a total load for the period of 2.46×10^6 mtons, an average of 223.5×10^3 mtons/day, 41.3% of the USGS estimated total load. On the basis of this analysis, it seems safe to conclude that the LWG models substantially underestimate sediment loads during flood periods.

The February 1996 flood

The flood of February 1996 was the second largest flood event of the last 60 years, but few data were collected during this event (Figure 17). As with the 1964 flood, the hysteresis models predict substantially higher transports than the LWG models, and it is more consistent with the data. Little effect of hysteresis is seen, but this may be an artifact – were 6-hr discharge data available, then hysteresis effects might be seen. Again, it appears to be important to represent the sediment input and river flow at intervals shorter than one day, and the LWG models under-predict sediment loading to Portland Harbor. Finally, the peak discharge in 1996 was likely larger than the 11,890 m^3/s daily average value for 9 February 1996, but it is unclear whether it was as large as the peak discharge in 1964, 12,540 m^3/s.

Annual Average Sediment Loads

Annual average total sediment loads for the 1962-2013 period are shown in Figure 18[2]. The annual average total load over the period from the hysteresis model is 1.85 X 10^6 mtons; that from the Appendix LA model is 1.11 X 10^6 mtons. It appears likely that the LWG model in Appendix La considerably underestimates the actual total load. These loads must be correct, or SEDZLJ and QEAFATE predictions will be incorrect.

[2] The loads here are in mtons, not short tons; the values in short tons would be about 10% higher. Also, the year-to-year variations shown in Figure 18 do not exactly track those in Figure 2-57 of Appendix La for reasons that are not clear. Part of the difference probably stems from the fact that loads are compiled here by water year rather than calendar year. The water year begins 1 October. Thus, water-year 1964 began 1 October 1963.

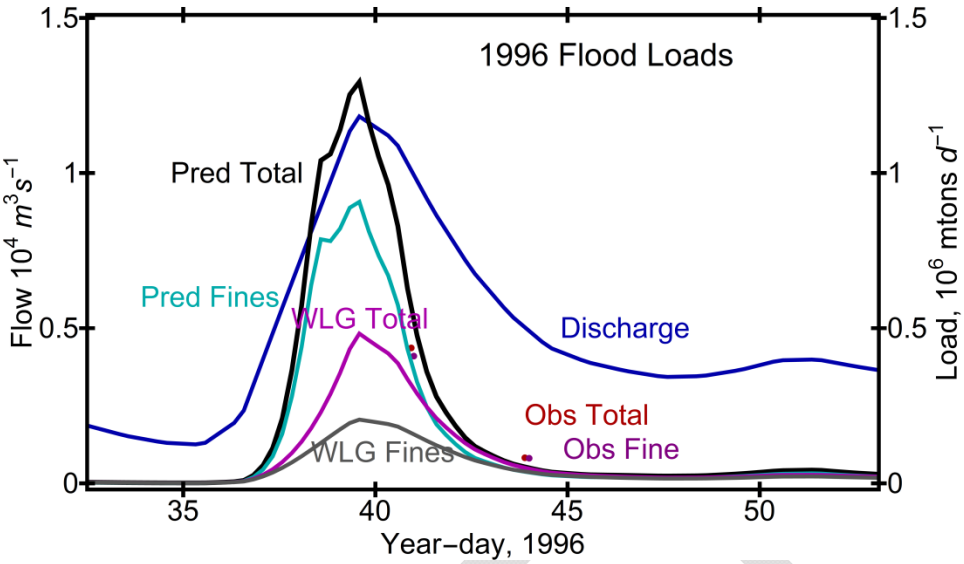

Figure 17: The observed and modeled total and fine loads for the 1996 flood. As in Figures 15 and 16, a ~21 day period is modeled. Observations are scant, and the predictions are based on daily data, because 6-hr discharge data are not available. The load is presented in 10^6 mtons/day. The fines and total load observations were collected at the same time, but the fines load data have been offset slightly in time for clarity.

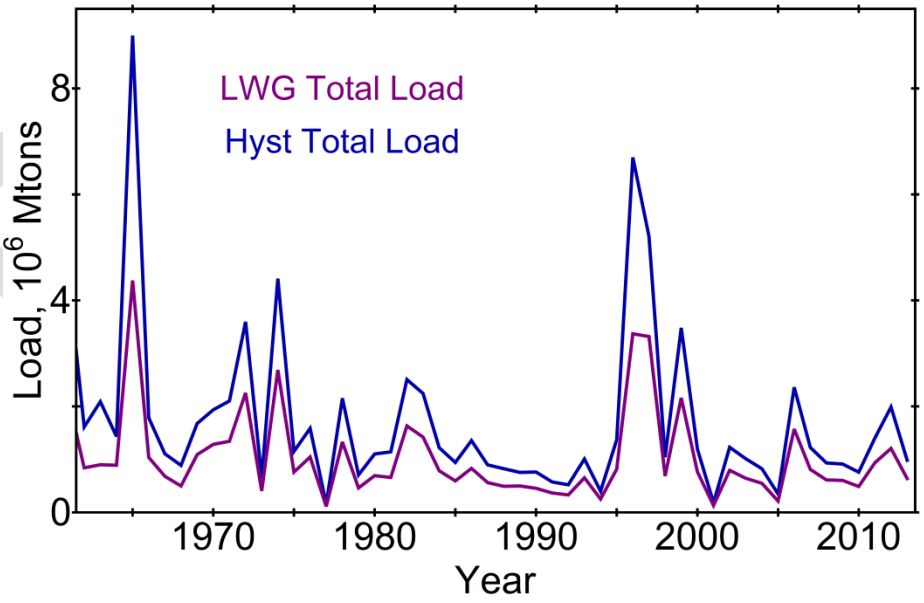

Figure 18: Annual average Willamette River total loads for water-years 1964 to 2013, based on the hysteresis model of (1) and (2) and the Appendix La model. The water-year begins 1 October.

Discussion

The above analysis of lower Willamette River sediment load suggests several issues:

Load model exponents: Sediment load typically increases with a power of discharge greater than 1. For high flow levels, the effective exponent in the hysteresis model represented by (1) is the sum of parameters $b + c$. This is 2.7 for the total load, 2.6 for fines load, and 3.57 for sand load. That the sand load has a higher exponent is typical. For comparison, the total load and fines load exponents in the WLG models are 2.65 for high flows for the total load, but the fines load does not increase nearly as rapidly at high flows, due to the percent fines formulation. That the exponent in the WLG total load model is lower probably reflects the lack of flood-event data used for calibration.

Use of a transport threshold Q_C: While a critical discharge Q_C is useful in representing the lower Willamette River sediment load, the meaning of Q_C is unclear. There are several possibilities; it might represent: a) sediment supply effects on the fine load that typically come into play at this discharge either from the mainstem or tributaries below Oregon City; b) a local increase in erosion of the bed at this discharge; c) the discharge at which the reaches above the Morrison Street Bridge change from being net deposition to net erosional; or d) the effects of the impoundment at Willamette Falls in Oregon City. These possibilities are not mutually exclusive, and more than one may be involved. Figure 2.6 of LWG [2006] suggests, however, that this change of slope persists at least as far landward as RM-17, rendering c) less likely.

Sediment load hysteresis: Hysteresis is quite prominent in the lower Willamette River sediment load. It presumable derives in part from supply limitation on fines input. Because rain-on-snow-events account for most Willamette River floods, it may also be related to snow melt.

Variability of the sediment load: The discharge and sediment load patterns during the 1964 flood strongly suggest that a single daily value does not adequately represent discharge or the sediment load estimated from the discharge. The peak discharge of 12,545 m^3/s on 25 December was more than 300 m^3/s greater than the daily average discharge for the day (12,233 m^3/s). Also, discharge increased from 2719 to 7787 m^3/s in a 24 hr period on 22-23 December. This caused an order of magnitude increase in instantaneous load.

The role of flood events: Large floods during the 1860 to 1920 period are thought to have had a major impact on evolution of the Willamette River channel above Oregon City [Wallick et al., 2007]. Little thought has been given to the importance of such events in Portland Harbor – are the configuration of the system and the contaminant distribution driven primarily by the annual flow cycle or by much less frequent, but very large events? The answer to this question (which is not known) has major implications for modeling strategy, because: a) it governs what sort of events should be simulated, and b) it has implications for evaluating model results – should all size classes be mobile at some times, or are some sediments not the product of modern processes?

Change in load properties: Several impoundments have gone online in the Willamette River Basin since the 1962 to 1964 data set was collected. There are no obvious differences between the 1962-1964 and

the post 1974 data sets, but we do not know whether load characteristics have changed over time, due either to impoundments or changes in land use.

The importance of reservoirs in the system: The reservoir capacity of the Willamette is <30% of the annual average. Both the 1964 and 1996 flood discharge maxima were considerably reduced by reservoir storage, which presumably also reduced sediment input. Both were also rain-on-snow events. A longer sequence of events, such as occurred in December 1922 to January 1923, could overwhelm the reservoir system, leading to a flood of a magnitude that has not occurred since the 19th century, despite the reservoir system. The 1922-1923 event is also noteworthy for the fact that snowmelt played a relatively small role in the flooding – the snowpack that had accumulated was gone well before the peak in the flow [Brands, 1947].

The foregoing discussion suggests that there are important unknowns in terms of process understanding, and critical weaknesses in representation of the sediment load at the upstream boundary of Portland Harbor. Providing adequate estimates of sediment load for modeling purposes will likely require both improved conceptual understanding of factors governing the load and better analytical tools for prediction of sediment loads.

The importance of the clay load: the table reproduced in Figure 19 from Waananen et al. [1970] emphasizes how large the clay load is during rain-on-snow events. Note that clays range from 33 to 61% of the total load, increasing as the flows decrease. Silts peak on the rising arm of the flood with the hysteresis and decrease thereafter. Sands peak just after the peak flow. This strongly suggests that clays need to represented as a size or settling class in numerical modeling, and that the sands, silts and clays loads all need to be defined separately, based on new observations.

Instantaneous suspended sediment and particle size, 1964-65

Date of collection	Hour	Water temp. (°F)	Discharge (cfs)	Specific conductance (micromhos at 25°C)	Suspended sediment				
					Concentration (ppm)	a Tons per day	Percent in size class indicated, in millimeters		
							Clay (less than .004)	Silt (.004-.062)	Sand (.062-1.00)
Dec. 23.....	1610	44	332,000	-	1,830	1,670,000	33	56	11
Dec. 25.....	1000	50	443,000	-	925	1,210,000	46	39	15
Dec. 26.....	0910	46	403,000	-	650	777,000	45	38	17
Dec. 29.....	1125	42	258,000	-	240	177,000	61	30	9

a Total sediment load.

Figure 19: Willamette River sediment load characteristics from samples collected at the Morrison Street Bridge during the December 1964 flood; from Waananen et al. [1970].

Summary and Conclusions

Correct representation of sediment input to Portland Harbor is vital to accurate modeling of erosion and deposition in the system and to the fate and transport of contaminants. It is important to represent not only the mean inputs of fines and total load, but also the timing – flood events will typically exhibit a sequence of erosion followed by deposition, and the depth of erosion is critical to contaminant

transport. If the sediment input is incorrect in timing or amount, then the sequence of events during a flood cannot be accurately modeled. Moreover, if sediment input is incorrectly estimated, as appears to be the case with the WLG models represented by (3) to (7), the sediment transport representation will likely have to be adjusted in unrealistic ways in order to avoid very unrealistic deposition and erosion patterns.

There are at least five difficulties in modeling lower Willamette River sediment load: a) the strong hysteresis of the load; b) the change in slope of the rating at a discharge of about 125% of the mean flow; c) the only daily data (1962-1964) are now 50 years old and may not represent present conditions; d) the sand silt and clay loads all need to be modeled separately because they behave differently, but data are not available to accomplish this; and e) the USGS load data were not collected at the upstream boundary of Portland Harbor and do not actually represent sediment input at the boundary of the system. The first two difficulties can be adequately dealt with in a load model like (1) and (2), but the last three require collection of new data. Further, the simple WLG models of (3) to (7) do not adequately represent the sediment load observations for the Willamette River at Portland. They underestimate mean load and strongly underestimate flood loads, while also incorrectly representing the timing of the load during flood events. The hysteresis model based on (1) and (2) has a modest tendency to overestimate the mean load and may strongly overestimate loads during small floods, though it reproduced the total and fines loads of the very large 1964 flood quite well. Importantly, it correctly times the peak sediment input, during the rising hydrograph. On the whole, it is considerably better (as well as more complex) than the WLG models.

From the forgoing, it should not be inferred that the hysteresis model is correct or authoritative. On the contrary – while it is both more effective than, and in considerably disagreement with, the WLG models, it remains a simplistic and imperfect tool. As noted by Meade et al. [1990] rating curve models of sediment load are at best approximate tools that are often incorrect. The best that can be said from the disagreement of these two approaches to estimating the sediment load is that sediment inputs to Portland Harbor remain poorly understood and poorly modeled. There are three steps that should be taken to improve this situation:

1. Sediment load should be determined on a continuous basis by multiple fixed instruments (at several depths) at a cross-section near the upstream boundary of Portland Harbor (not the Morrison Street Bridge) for several years, to capture a wide range of conditions. Given that fines predominate in the load, turbidity time series should provide good estimates of sediment load; cf. USGS [2009]. For calibration, the present NWIS-style monthly sampling would need to be augmented by multiple samples during major flood periods. Given that the clay load is prominent, at least during floods, a more refined analysis of sediment load sizing than sand vs. fines is needed, and numerous size samples will be needed to calibrate the turbidity.

2. The Lower Willamette River sediment load needs to be examined on a seasonal basis, to determine whether how hysteresis effects vary seasonally. This exercise can be begun with the existing data set and would help provide a more mechanistic understanding of Portland

Harbor sediment loading.

3. A watershed modeling approach to estimating sediment input should be developed. While such a model would need to be calibrated, there are now many approaches to this sort of modeling [Borah et al., 2008].

In conclusion, sediment load modeling is a major weakness in the LWG Portland Harbor Superfund site modeling efforts. Loads are not correctly represented by the rating curves employed, with consequences for numerical model accuracy that should be assessed, but have not been.

References

Borah, D., Krug, E., and Yoder, D., 2008. Watershed Sediment Yield. Sedimentation Engineering: pp. 827-858. doi: 10.1061/9780784408148.ch17.

Brands, M .D., 1947. Flood runoff in the Willamette Valley, Oregon, USGS Water Supply Paper: 968-A, Washington, D. C.

Duan, N., 1983. Smearing estimate – a non-parametric retransformation method. *Journal of the American Statistical Association,* 78(383), 605-610.

EPA Region 2 2014. Focused Feasibility Study Lower Eight Miles of the Lower Passaic River, Appendix B2 Lower Passaic River Sediment Transport Model, EPA, New York.

Gray, J.R., and Simões, F.J.M., 2008. Estimating sediment discharge, Appendix D in Garcia, M.H., ed., *Sedimentation engineering—processes, measurements, modeling, and practice*: American Society of Civil Engineers Manuals and Reports on Engineering Practice No. 110, p. 1067–1088.

Haushild, W. L., R. W. Perkins, H. H. Stevens, G. R. Dempster, and J. L. Glenn, 1966. *Progress report: radionuclide transport in the Pasco to Vancouver, Washington reach of the Columbia River July 1962 to September 1963*. Portland, Oregon, U. S. Geological Survey. 188 p. (Prepared in co-operation with the U. S. Atomic Energy Commission.)

Lower Willamette Group (LWG), 2006. Portland Harbor RI/FS Phase 2 Recalibration Results: Hydrodynamic Sedimentation Modeling for the Lower Willamette River Draft, prepared by WEST Consultants and Integral Consulting, Inc.

Meade, R. H., T. R. Yuzyk, and T. J. Day, 1990. Movement and storage of sediment in rivers of the United States and Canada. Pages 255–280. in Wolman, M. G. and H. C. Riggs, editors. eds. Surface Water Hydrology, vol. 0–1: The Geology of North America. Boulder (CO) Geological Society of America.

USGS, 2009. Guidelines and Procedures for Computing Time-Series Suspended-Sediment Concentrations and Loads from In-Stream Turbidity-Sensor and Streamflow Data, Chapter 4 of Book 3, Applications of Hydraulics Section C, Sediment and Erosion Techniques. Techniques and Methods 3–C4, US Department of Interior, USG Geological Survey, Reston, VA.

Waananen, A.O.; Harris, D.D.; Williams, R.C., 1970. Floods of December 1964 and January 1965 in the Far Western States; Part 2 Streamflow and Sediment Data. 1971, USGS Water Supply Paper: 1866-A, Washington, DC.

Waananen, A.O.; Harris, D.D.; Williams, R.C., 1971. Floods of December 1964 and January 1965 in the Far Western States; Part 1 Description. 1971, USGS Water Supply Paper: 1866-A, Washington, DC.

Wallick, J.R., Grant, G.E., Lancaster, S.T., Bolte, J.P., and Denlinger, R.P., 2007. Patterns and controls on historical channel change in the Willamette River, Oregon, *in* Gupta, A.V., ed., Large Rivers: Geomorphology and Management, John Wiley and Sons, Chichester, p. 491–516.